

附件1

江苏省高质量数据集重点领域首批建设清单

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
1	工业制造	汇聚生产过程、质量控制、检定标准、物料与供应链等关键基础数据，面向研发设计、中试验证、测试评价、生产制造、经营管理、产业链供应链协同、运维服务等典型场景，分行业打造高质量数据集	省工业和信息化厅、省市场监管局	南京、无锡、常州、苏州、南通、连云港、淮安	钢铁行业生产、销售、采购、成本管理数据集	焦球团、炼铁、炼钢等生产线的研发、工艺、制造、管理、销售、成本的数据集	智能质量诊断、工艺过程修正、服务预测、研发产品优化
				南京、常州、镇江	电气行业多模态高质量数据集	电气行业设计、供应链、制造、运维等产品全生命周期关键环节的数据集	智能设计辅助、制造过程自适应优化和设备智能预测性维护
				无锡、徐州、苏州、南通、扬州	高端装备行业高质量数据集	高端装备制造、研发、供应链、售后服务等环节的数据集	智能制造优化、设备远程监控与维护、产品性能提升
				南京、苏州、扬州	智能电网行业高质量数据集	智能电网调度、运行、维护、用户服务等环节的数据集	电网智能调度、故障预测与定位、能源管理优化
				南通	家纺行业高质量数据集	家纺行业设计画稿、历史花型、工艺细节、生产效能、订单履约等数据集	提高物流配送效率与可视化程度，减少退换货成本，原材料溯源
				扬州	日化行业高质量数据集	聚焦清洁用品、个人护理等，整合生产工艺、材质特性、供应链协同等日化行业数据集	支撑高端产品创新、牙膏、防晒霜等日化产品供应链与渠道优化、精准营销与消费者洞察
				南京、无锡、苏州	具身智能机器人行业高质量数据集	具身智能机器人模型训练与优化、应用部署等环节的数据集	具身智能机器人在汽车、3C等行业应用推广

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
2	交通运输	汇聚交通规划、交通流量等关键基础数据，打造面向交通智能调度管理、铁路智能运营和维护等典型应用场景的高质量数据集	省交通运输厅	南京、无锡、淮安、宿迁	公路货运与交通信号数据集	公路货运订单交易及城市交通信号控制的数据集	大模型训练优化、实时交通信号优化、路况分析与预测
				南京、无锡、常州、苏州、淮安	自动驾驶行业高质量数据集	路端感知、特定场景与任务、4D成像雷达等数据集	自动驾驶模型训练、智能交通管理、交通规划支持
				南京	网约汽车司乘服务数据集	司乘服务、智能调度、出行保障、订单信息、行程录音等数据集	司乘服务、智能调度、客服接单、路线规划、算法优化等应用
				南通、连云港、淮安、泰州	港航数据高质量数据集	船舶、货物、物流等数据集	物流跟踪、电子单证、业务预报等应用
3	医疗健康	汇聚医疗服务、医院管理、药品配备使用、安全用药、公共卫生、健康产业、医保结算、医药耗材、医学教学科研、中医药健康服务、疫情防控等关键基础数据，打造面向临床辅助决策、医学影像诊断、医院治理、医保基金监管、公共卫生、基层健康管理、医用机器人、药物研发、药物警戒、药物疗效评估、传染病防治、医药流通、商保服务等应用场景的高质量数据集	省卫生健康委、省医保局、省工业和信息化厅	南京、无锡、徐州、常州、苏州、连云港、盐城	医学高质量数据集	手术医疗信息、治疗反应、疾病进展等医学数据集	医疗服务优化、多模态辅助诊断、人工智能辅助诊疗、疾病分级分型、术前风险评估、术中导航、患者预后评估
				南京、苏州、南通、连云港、盐城、扬州、泰州	卫生健康信息数据集	医疗门诊、住院、检验检查、影像、药品及耗材使用、健康档案管理、慢性病管理、老年保健、计划免疫服务等医疗和公共卫生数据集	卫生健康分析及决策支持，疾病监测及预防控制，临床辅助决策支持，公共卫生应急响应，医疗质量监控
				省级、南京、苏州、连云港	医保结算与药品追溯数据集	全省医保定点医院、医保定点药房的医保费用结算、药品耗材交易、诊疗就診登记的数据集	医保基金智能监管、医药产业发展、商保医保服务、医药零售服务

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
				常州、苏州、连云港、宿迁	生物信息学高质量数据集	肿瘤、血液病、代谢病、免疫病等数据集	新药研发、临床诊疗、基础医学研究等AI大模型训练
4	科学研究	汇聚通过实验研究、试验开发、观测监测、资源调查、检验检测、计量测量等方式产生并用于科学研究活动的关键数据，打造面向基础研究、应用研究、技术创新、成果转化与产业化等典型应用场景的高质量数据集	省科技厅、省知识产权局	无锡、扬州	生物大分子高质量数据集	基因组学DNA、转录组学RNA、生命基座蛋白质等生命体必备大分子序列、功能、结构数据集	生命体必备大分子功能、结构、交互性和研发性智能理解和设计，大分子药物研发筛选
				苏州	全球公开专利数据集	网页、百科、书籍、试题数据库、代码仓储等，以及全球专利、诉讼记录等公开数据形成的数据集	知识产权行业大模型
				无锡、苏州	新材料高质量数据集	创新重构性材料体系研发范式的数据集	先进半导体、新能源、军工等战略产业应用
				无锡	芯片设计领域高质量数据集	用于芯片设计的代码验证和优化的数据集	芯片设计、EDA工具开发、人工智能在芯片设计中的应用
				无锡	电子元器件建模模型高质量数据集	先进工艺节点输出数据、转移数据、电容等数据集	器件建模开发以及电路仿真设计
5	金融服务	汇聚金融管理、金融业务等关键基础数据，打造面向金融风险防控、保险产品设计、精准营销等典型应用场景的高质量数据	人民银行江苏省分行、省国资委	省级、南京、苏州、南通、盐城、宿迁	企业信用数据集	民营企业的基本信息、财务指标、经营状况、守法合规、关联关系、信贷记录、行业趋势等多维度海量涉企数据	金融机构风险控制

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
		集				的数据集	
6	文化旅游	汇聚文物、古籍、近现代文献、地方志、非物质文化遗产、民族民间文艺等关键基础数据，打造面向文化创意旅游、展览、文物保护修复、文化传承、阐释传播等典型应用场景的高质量数据集	省文化和旅游厅、省地方志工作办公室	南京、无锡、苏州、淮安、扬州、宿迁	文旅行业垂直数据集	文旅相关数据的行业垂直数据集	数据驱动智能决策，洞察分析市场行情
				南京、苏州	丝绸纹样数据集	标本库、素材库、基因库等丝绸纹样数据集	丝绸纹样版权合作
				扬州	民族乐器数据集	曲谱数据、音频标注数据、演奏数据、乐器状态监测数据、乐器品质检测数据、用户行为数据等数据集	乐器生产厂商、乐器研究机构、乐器使用商家、艺术培训机构、教育机构、音乐家庭等应用
				省级	文化传媒行业高质量数据集	汇聚文化传媒行业文本、图像、音频、视频、3D/VR和结构化等模态数据，涵盖报纸文献、航拍影像、广播剧、虚拟场景、舆情评论和多语言新闻等丰富内容的数据集	红色教育大模型微调、二次创作、长篇摘要、多语言机器翻译、纪录片智能创作、虚拟主播

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
				省级	地方志高质量数据集	全省各级志书、年鉴中的文字、图片和表格等通用数据集和经济、文化、社会等特定领域垂直数据集	经济发展、社会治理、文化遗产、政务服务等领域大模型训练优化
7	城市治理	汇聚自然资源调查监测评价、国土空间规划、测绘基准、实景三维、地理信息、不动产登记、房屋建筑、市政设施、人口统计、交通流量、环境质量、经济指标、工程建设、城市信息模型等关键基础数据，打造面向城市建设运行、公共卫生、交通管理、生态环境、基层治理、住房公积金服务等典型应用场景的高质量数据集	省住房城乡建设厅、省自然资源厅、省交通运输厅、省生态环境厅、省公安厅	南京、无锡、常州、苏州、盐城、扬州、镇江、宿迁	低空经济影像数据集	针对雷达点云、Mesh倾斜三维模型、航摄影像、真彩视频、红外视频等多模态数据，构建地物、动态目标的多模态标签识别数据集、地名地址信息数据集	危险建筑识别、路况预测分析等城市精细化治理
				无锡、常州、苏州、南通、淮安、盐城、扬州、泰州、宿迁	城市治理数据集	基础地图与地理编码数据、城市部件与事件监管数据、执法管理数据、环卫管理数据、共享单车管理数据、广告与市容数据、执法资源数据的数据集	城市治理全流程闭环，智能分析与决策，跨部门协同治理，公众服务优化

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
8	人力资源	汇聚人员信息、招聘选拔、薪酬福利、绩效考核、职业发展、职业资格、社保缴纳等关键基础数据，打造人力资源服务、人才培养使用、智能监管服务等典型应用场景的高质量数据集	省人力资源社会保障厅、省委组织部	省级、泰州	人力资源社会保障行业数据集	个人社保缴纳信息、企业社保缴纳信息、职业资格证书信息、劳动仲裁信息、劳动监察信息的数据集	赋能公共服务精准惠民、信用评价
9	绿色低碳	汇聚碳排放、固体废物收集、转移、利用、处置各环节、生态环境、土地、矿产、海洋、遥感测绘等关键基础数据，打造能源清洁高效利用、温室气体排放控制、污水处理、自然资源和地理空间、生态修复等典型应用场景的高质量数据集	省发展改革委、省生态环境厅、省自然资源厅	省级、南京、无锡、苏州、南通、扬州	碳排放监测数据集	企业碳足迹数据、能源消耗数据、碳交易记录、清洁能源使用数据等数据集	碳排放监测、能源管理优化、绿色低碳政策制定与评估
				扬州	污水处理全流程运营分析数据集	设备运行、水质监测等数据集	节能减排、决策支持
10	农业农村	汇聚农业农村关键基础数据打造面向智慧种养、智慧监管、智能农机应用和乡村发展等典型应用场景的高质量数据集	省农业农村厅	南京、无锡、常州、苏州、淮安、盐城、扬州、宿迁	农业生产信息数据集	土壤墒情数据、病虫害预警数据、农产品批发市场价格数据等数据集，面向农机驾驶的机耕道路网和农田边界信息等数据集	农作物病虫害预警、农机作业监测、农情监测

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
11	智慧能源	汇聚能源资源、能源生产与供给、能源消费、能源基础设施等关键基础数据，打造面向智慧能源管理、清洁能源与储能应用、能源交易等典型应用场景的高质量数据集	省能源局	无锡、镇江、宿迁	能源管理数据集	能源消耗监测、电池性能分析、需求响应管理等数据集	能源健康管理
12	教育教学	汇聚课程与教学、学科与专业、教育设施、教学质量评估等关键基础数据，打造面向智能教学、互动课堂、个性化学习等典型应用场景的高质量数据集	省教育厅	南京、连云港	高职院校课程数据集	课程名称、学分、适用专业等基本信息的结构化课程数据集，教学视频、课件文档、虚拟仿真实验及实训案例的数字化课程资源数据集	产教融合、人才培养
				南京、无锡、苏州、南通、连云港、扬州	教育行业场景化AI训练数据集	学生课堂行为、作业数据、校园安全预警、校园监控视频、考勤记录等数据的教育行业场景数据集	校园安全、辅助学习

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
13	商务领域	汇聚消费、流通、贸易、投资、广告营销等领域的关键基础数据，打造面向新型消费、智慧商圈、数字供应链、数字贸易、广告产业分析、创意设计、品牌营销、公益宣传等典型应用场景的数字商务和广告营销高质量数据集	省商务厅、省市场监管局	南京、无锡、苏州、淮安、盐城、泰州、宿迁	商务行业数据集	经营主体的注册登记数据、司法风险、知识产权、经营信息、招投标、经营风险、上市公司、企业发展、特殊主体等数据集	金融机构、政府监管、企业服务等
				无锡、南通、宿迁	电商交易情况数据集	电商交易相关的企业、店铺、品类等数据集	区域电商交易分析、大屏展示
14	应急管理	汇聚危险化学品、矿山、工贸等行业安全生产和地震、洪涝、林火等自然灾害关键基础数据，打造面向监测预警、监管执法、抢险救援等典型应用场景的高质量数据集	省应急管理厅	南京、无锡、徐州、苏州	地质行业数据集	卫星遥感影像数据、海洋监测数据、森林资源数据等数据集	自然资源全要素调查监测、海洋监测预警、耕地保护、生态修复、地面沉降监测
				徐州	煤矿多灾害融合监测数据集	瓦斯数据、环境数据、顶板与围岩数据、机电设备运行数据、人员定位数据、视频监控数据等数据集	矿山安全预测预警，服务于煤矿、金属和非金属矿山瓦斯、矿压、冲击地压、突水和煤自燃等灾害的监测预警
15	气象服务	汇聚大气、陆地、海洋等地球系统多圈层的现场观测、遥感遥测、预报预测等关键基础数据，打造面向气象防灾减灾、气候变化、气象服务等典型应	省气象局	南京、无锡、宿迁	气象行业数据集	高精度气象观测数据、极端天气预警数据、气候模型预测数据、实时观测数据、空气质量监测数据、长期历史气象观测数据、历史极端天气等数据集	气象预报服务、气象灾害预警、气候研究与评估、长期天气模式预测

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
		用场景的高质量数据集					
16	公共安全	汇聚公共安全风险预警、预防打击违法犯罪、社会治理与服务民生、视频监控、指挥调度、自然人信用信息等关键基础数据，打造面向智慧安防、“情指行”、侦查办案、智慧警务、自然人信用应用等典型应用场景的高质量数据集	省公安厅、省应急管理厅、省交通运输厅	南京、淮安	警情风险感知模型数据集	涉警地址地理空间数据集，涉警家庭数据集，接警信息文本和出警信息文本数据集，警情关联风险数据集，警情风险程度数据集	警情分析
				常州、苏州、淮安	城市公共场所异常行为监测数据集	行人、车辆正常与异常行为、危险品车辆电子运单报备、危险货物港口作业申请、危化品生产企业清单及位置分布等数据集	智能安防、危化品管理
17	其他	除上述重点领域外，通用大模型、数据跨境、政务服务等其他领域的高质量数据集	省委网信办、省数据局	无锡、苏州、淮安	跨领域多模态语音交互数据集	智能家居、车载系统、客服对话等多个领域场景，语音、文本、图像等多模态数据的跨领域多模态语音交互数据集	多模态通用人工智能大模型
				南京	电池行业数据跨境数据集	以产品数字护照为载体，电池全生命周期、ESG信息等跨境数据集	数据跨境合规流动、供应链优化升级
				苏州	网络安全行业通用安全知识数据集	围绕网络安全框架IPDRR展开，涵盖识别、保护、检测、响应和恢复等环节，具有风险评估、访问控制、事	安全知识问答、网络安全运营、网络安全处置建议、网络安全解决方案生成等应用场景

序号	场景领域	建设方式	主管部门	牵头建设地市	数据集名称	建设内容	典型应用场景
						件响应、漏洞修复、威胁检测等核心内容的数据集	
				省级、南京、苏州、淮安、扬州、镇江	12345热线高质量数据集	基本信息、诉求详情、诉求处理过程、诉求处理结果、服务质量评价等数据集	智能派单、紧急事件识别、政策效果追踪、数据开放服务等应用
				省级、苏州、南通、淮安	公共资源交易数据集	招标、投标、评标等公共资源全流程交易信息、主体信息、信用信息、监管信息等数据集	招标文件合规检测、围串标识别、智能评审、智能交易服务等应用